

Prediction of equilibrium water dew point of natural gas in TEG dehydration systems using Bayesian Feedforward Artificial Neural Network (FANN)

Ahmad Z^{a,*}, Alireza Bahadori^b, Jie Zhang^c

^{a,*}School of Chemical Engineering, University Sains Malaysia, Engineering Campus, Seri Ampangan, 14300, Nibong Tebal, Penang, Malaysia

E-mail: chzahmad@usm.my

^bSchool of Environment, Science and Engineering, Southern Cross University, Lismore NSW Australia.

Email: Alireza.Bahadori@scu.edu.au

^cSchool of Chemical Engineering and Advanced Materials, Newcastle University, Newcastle upon Tyne NE1 7RU, UK

E-mail: jie.zhang@newcastle.ac.uk

Abstract

The aim of this paper is to predict the equilibrium water dew point of natural gas in TEG dehydration process using feedforward artificial neural network (FANN) and validation with the literatures values for generalization capability of the model. The FANN was trained by the Bayesian with the regularization method. The FANN consists of three layers which are input, hidden and output layer. The input layer is the manipulated variables which are contactor temperature in Kelvin and the TEG concentration in percent while the output layer is the equilibrium water dew point of the natural gas. The performances of the model were evaluated using the coefficient of determination, mean square error and average absolute deviation percentage. The predictive model shows a good result as the coefficient of determination of 0.9989 and 0.9976 was obtained for training and testing data respectively. Small value of mean square error of 0.0203 and 0.0221 was obtained for the training and testing data as well. 0.5 % of average absolute deviation percentage was observed which is comparable with the literatures. It clearly shows that FANN gives a good prediction on water dew point of natural gas in TEG dehydration process even though the process is highly dynamics.

Keywords: process modelling, TEG dehydration process, Equilibrium water dew point, artificial intelligent, feedforward artificial neural network

Introduction

Artificial neural network (ANN) had emerged as an attractive tool for non-linear multivariate modelling in the last two decades (Desai et al., 2008; Zainal et al., 2010). It had typically been used as a “black-box” tool, which was, no prior knowledge about the process was assumed but the goal was to develop a process model based only on observations of its input-output behaviour (Psichogios and Ungar, 1992). In the other words, ANN was a mathematical system that simulated biological neural networks and was often described as a massively interconnected network structure consisting of many simple processing elements (neurons) with the ability to perform parallel computation for data processing (Agatonovic-Kustrin et

al., 1998; Baş and Boyacı, 2007). ANN was capable of handling multiple independent and dependent variables simultaneously and to do this prior knowledge on the functional relationship did not need to be known. Each neuron received information through input connections, processed the information and produced the output which was distributed through output connections. Each connection was associated with a real number quantity known as weight (Desai et al., 2008). In other words, the neurons summed weighted inputs and then applied a linear or non-linear function to the resulting sum to determine the output. These neurons were arranged in layers and were combined through excessive connectivity (Shen et al., 2007). A neural network in its basic form was usually composed of several layers of neurons, there being one input layer, one output layer and at least one hidden layer (Figure 1).

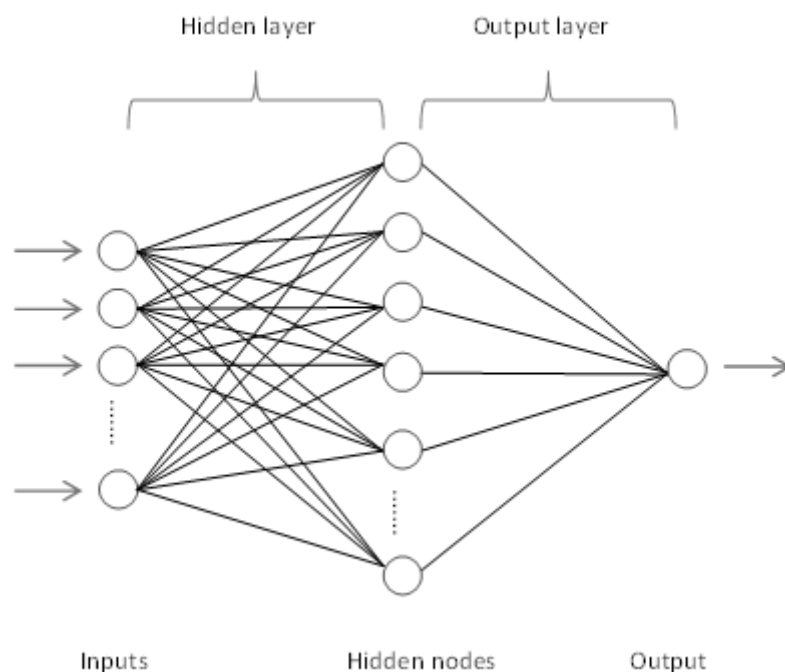


Figure 1: Generic architecture of ANN. The lines connecting the neurons represent the weight.

There were various types of ANNs. Amongst them were the Hopfield, Kohonen, Radial Basis Function and the most common and popular network was the feedforward artificial neural network (FANN). FANN was normally applied in analytical applications and was very powerful in function optimization modelling (Lou and Nakai, 2001). Information from various sets of inputs was fed forward through the network to optimize the weight between neurons, or to train it. As described by Agatonovic-Kustrin et al. (1998), the error or bias in prediction was then propagated through the system and the inter-unit connections were changed to minimize the error in prediction. This was a continuous process with multiple training sets until the minimum error was attained. Training an ANN network was

therefore accomplished by adjusting these weights by minimizing a non-linear error function (Desai et al., 2008).

The growing interests in applying FANN are due to the computing system that growth rapidly which enable the behaviour of the complex system to be modelled and predicted accurately. Furthermore the characteristic of neural network models them self that learn from examples rather than having to programme the complex system also contributed the application of the models. The architecture of FANN vary from multilayer perceptron to radial basis function (RBF) and also recurrent neural networks models (e.g. Hagan *et al.*, 2002). The Bayesian learning framework was first applied to neural networks by Buntine and Weigend (1992), MacKay (1992) and Radford (1997). There are several reasons that can explain on how the Bayesian learning rule increases the robustness of neural networks. First, instead of finding an estimate for the mean prediction of the models, it obtains an estimate for the entire distribution of model prediction. This estimate, which is take into account both the noise in the data and the variance of the models, which is valuable to evaluate the confidence for the individual prediction (Lampinen and Vehtari, 2001).

Most of the applications of FANN are concentrated on the modelling and control of chemical processes using multilayer perceptron networks. The common systems used in the chemical processes are distillation columns, and reactor systems (continuous stirred tank reactor (CSTR), bioreactor, and neutralising reactor). These processes are usually very nonlinear and nonlinear models have to be developed. Currently, applications of single neural networks in process modeling and control are quite significant in industry especially in model based predictive control (MBPC) (e.g. Chen and Yea, 2002; Xiong and Jutan, 2002) and this is due to the ability of neural networks in modeling nonlinear processes (e.g. Shaw *et al.*, 1997).

This paper is organised as follows. Section 2 presents the case study or the process. The concept of feedforward artificial neural network (FANN) modelling was presented in section 3. The result and discussion of the proposed modelling are presented in section 4. Finally, the last section concludes this paper.

Case Study: Triethylene Glycol (TEG) Process

The case study for the triethylene glycol (TEG) is taken from Bahadori and Vuthaluru (2009). TEG is the most common for natural gas dehydration and it is used in a counter-current mass transfer operation inside a contractor to establish the required water content in the outlet gas as shown in Figure 2. All the details of the process can be found in Bahadori and Vuthaluru (2009).

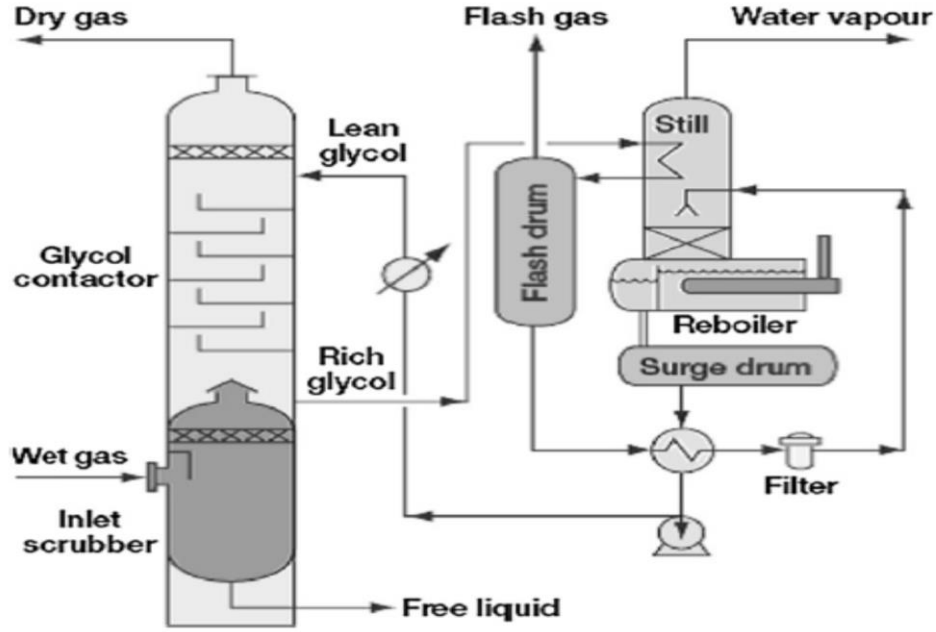


Figure 2: Typical TEG process (Bahadori, 2007)

Process Modelling

The required data to develop the prediction model includes the reported data (Parrish et al., 1986; Herskowitz and Gottlieb, 1984) for the rapid estimation of water dew point (T_d) of a natural gas stream in equilibrium with a triethylene glycol (TEG) solution at various contactor temperatures (T) and TEG concentrations (W) in percent. In this work, water dew point (T_d) of a natural gas stream in equilibrium with a triethylene glycol (TEG) solution is predicted using feedforward artificial neural network (FANN) trained by the Bayesian training algorithm.

The static model for water dew point prediction is of the form:

$$\hat{y}(t) = f[u_1(t), u_2(t)] \quad (1)$$

where \hat{y} represents the predicted water dew point in Kelvin, u_1 represent contactor temperature in Kelvin and u_2 represents the TEG concentration in percent, %, f is a nonlinear function represented by the neural networks and t is the discrete time. The procedures of constructing a FANN were summarized in Figure 2 which supported by Coit et.al. (1998) where in order to assure confidence in a black box model, the model must be constructed and validated using data which adequately and accurately reflected the process domain. Thus three types of data sets were used: (1) training data: used to train network; (2) testing data: used to monitor the neural network performance during training; and (3) validation data or unseen data: used to measure the performance of a trained network (Agatonovic-Kustrin et al., 1998). The number of the layers and processing elements in layers varied from one process to another. There was no strict rule available to identify the optimal number of layers

and neurons. In fact, it was generally defined by trial and error (Baş and Boyacı, 2007). Therefore for this case study, only 1 hidden layer is applied which is able to capture the dynamics of the process. In this study, all the data was normalized to zero mean with standard deviation to cope with the different magnitudes in the input and output data. Then, the input data were divided randomly using MatlabTM command *divideint* into three set of data which are 60% (99 samples) training, 26% (243 sample) testing and the remaining 10 samples as validation/unseen as show in Figure 3 and Table 1 after it has been arrange in discrete form as shown in equation (1).

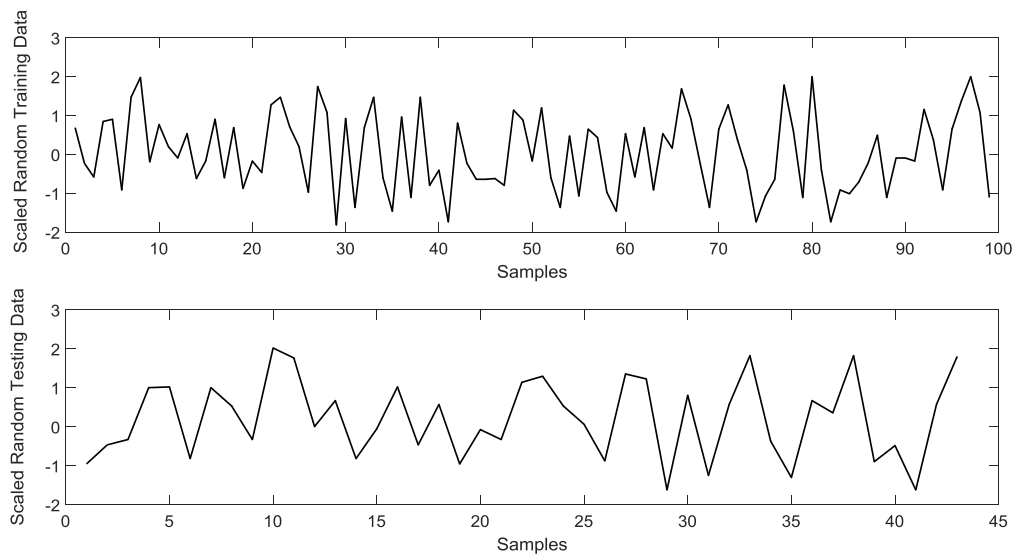


Figure 3: Scaled random data for Training and testing for process modeling.

Table 1: Unseen data for generalisation capability of the FANN model

TEG weight percent, %	Temperature, K	Actual dew point from Reported data (Gas Processors and Suppliers Association Engineering Data Book, 2004; Parrish et al., 1986; Herskowitz and Gottlieb, 1984), K
90	283.15	267.15
95	303.15	278.15
97	318.15	284.65
98	333.15	290.65
99	283.15	243.15
99.9	303.15	232.65
99.97	283.15	210.00
99.99	303.15	211.70
99.997	323.15	211.15
99.999	343.15	211.00

Then the neural networks were trained by the Bayesian algorithm. All weights and biases were randomly initialized in the range from -0.1 to 0.1 . In cross validation techniques neural network building data were divided into testing and training data sets and the mean squared error (MSE) for both were plotted against training epoch and when the testing MSE started to rise, the training was terminated and the value of weight that the network had in that previous step was chosen as the result of the training run.

Hidden neurons were associated with the logarithmic sigmoid activation function whereas output layer neurons used the linear activation function. To cope with different magnitudes in the input and output data, all the data were scaled to zero mean and unit standard deviation. The numbers of hidden neurons were determined through cross validation technique. Single hidden layer neural networks with different numbers of hidden neurons were trained on the training data and tested on the testing data. The network with the lowest mean squared errors (MSE) on the testing data was considered as having the best network topology. In assessing the developed models, MSE and correlation coefficient (R^2) on the unseen validation data were used as the performance criterion. The overall criterion for the FANN modelling is shown in Table 2.

Table 2: FANN network parameters

Parameters	Value
Epochs	1000
Target error goal	10^{-5}
Minimum performance gradient	10^{-5}
Number of hidden neuron	1 to 11
Number of hidden layer	1
Transfer function (Hidden layer)	Log-sigmoid (<i>logsig</i>)
Transfer function (output layer)	Linear (<i>purelin</i>)
Training algorithm	Bayesian Regularisation(<i>trainbr</i>)

Evaluation of Model Predictability

The performance of the FANNs was statistically measured by mean squared error (MSE) and coefficient determination (R^2) obtained as follows (Wang et al., 2008):

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - y_{di})^2 \quad (2)$$

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - y_{di})^2}{\sum_{i=1}^N (y_{di} - y_m)^2} \quad (3)$$

where N was the number of data, y_i was the predicted value by ANN model, y_{di} the actual or experimental value, and y_m was the average of actual values. The network having minimum MSE and maximum R^2 was selected as the best FANN model.

Result and discussion

Since there was no theoretical principle in choosing the proper network topology, several structures were tested to obtain the best one using cross validation technique. The numbers of nodes/neuron in the hidden layer were varied from 1 to 11 and the network was trained and tested after each addition of nodes/neuron. It was found that using 9 hidden nodes/neurons gives the least MSE on the training and testing data as shown in Figure 4. It clearly shows that increasing the number of nodes in the hidden layer will not significantly improve the model performance after 9 nodes. The final architecture of the FANN model is a 2 – 9 – 1 which represent two input nodes, 9 hidden nodes and 1 output nodes.

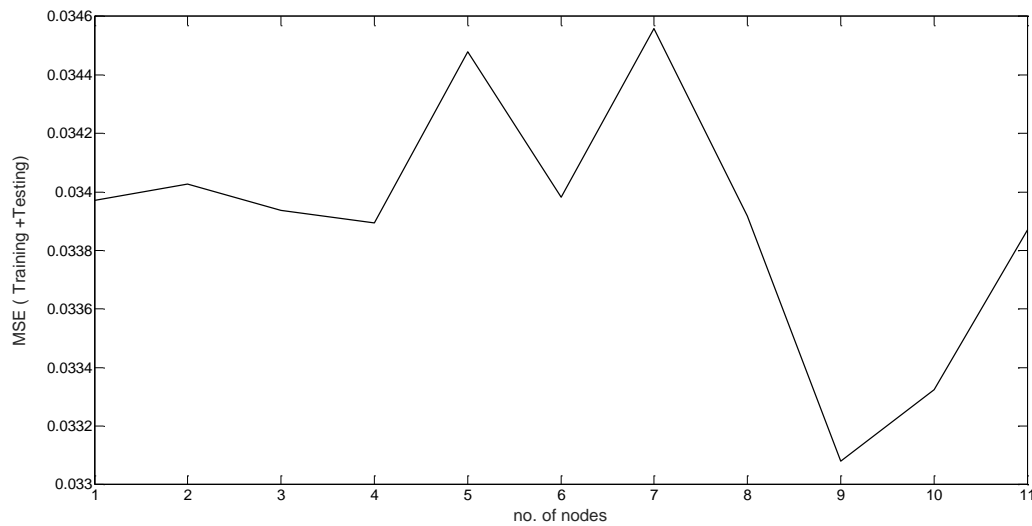


Figure 4: Number of nodes in the hidden layer corresponding to the MSE (Training and testing data).

The tan-sigmoid transfer function was used as the activation function for hidden layer, and linear transfer function was applied for the output layer. The performance of the modelling during training is shown in Figure 6. It clearly shows that the model able to reach the threshold setting while training. The predicted data for training and testing sets are illustrated in Figure 6. It was obvious that the actual and predicted values were more or less the same and in agreement with each other for the model with the 0.9988 and 0.9982 for R^2 for training and testing data respectively. The MSE for training and testing data are 0.0203 and 0.0221 respectively. The prediction for training and testing data is all in agreement and it shows that FANN with Bayesian training is able to capture the dynamics of the process of the equilibrium water dew point of natural gas in TEG dehydration system accurately. This model performance was supported by the regression and residue analysis in Figure 7 and 8 respectively. In the residue analysis, most of the samples errors are skewed toward the centred of zero error. It shows that the model developed is able to predict the process accurately.

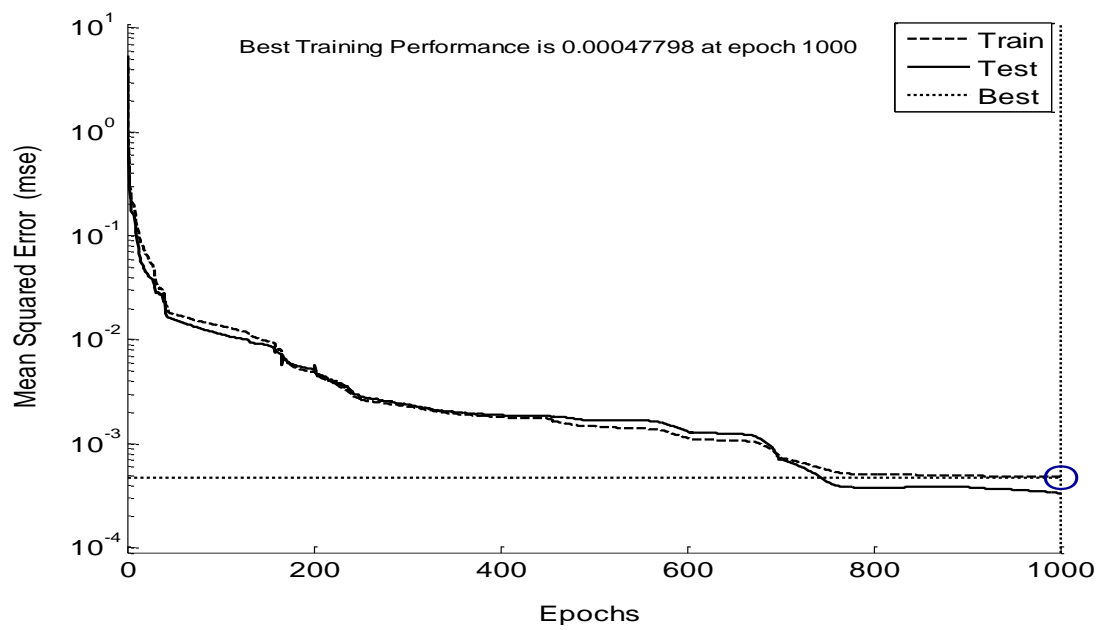


Figure 5: Epoch for the training profile

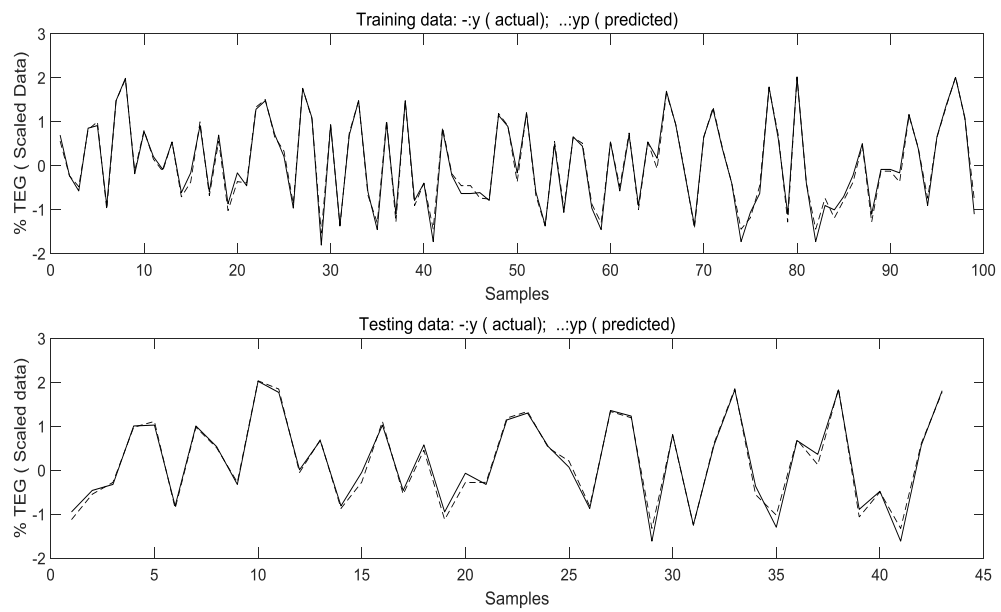


Figure 6: Training and testing prediction data

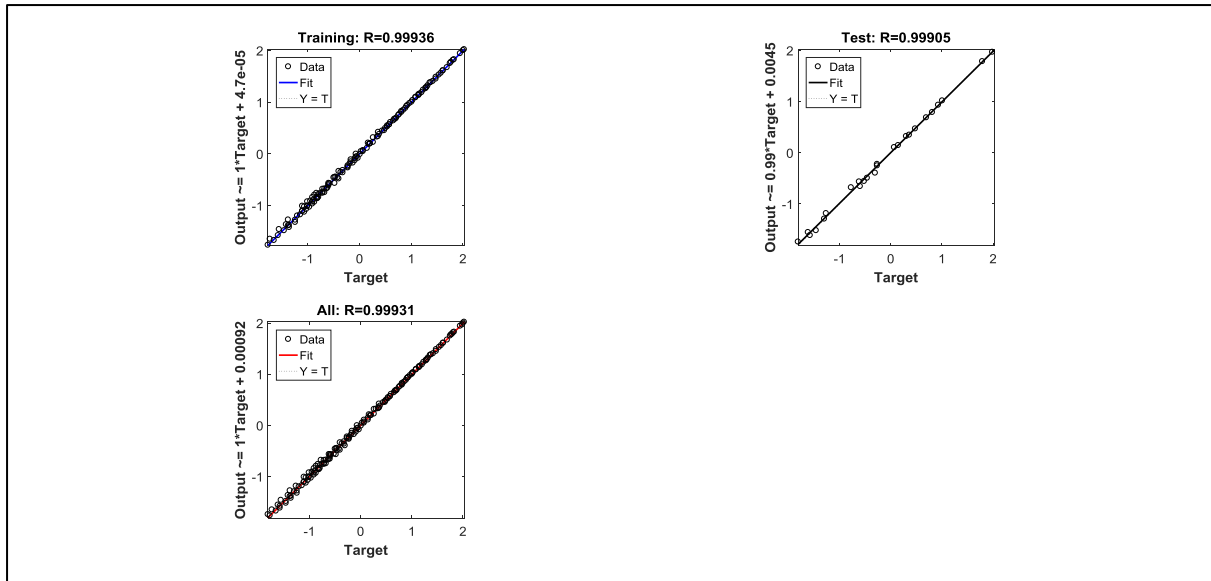


Figure 7: Regression for training and testing data.

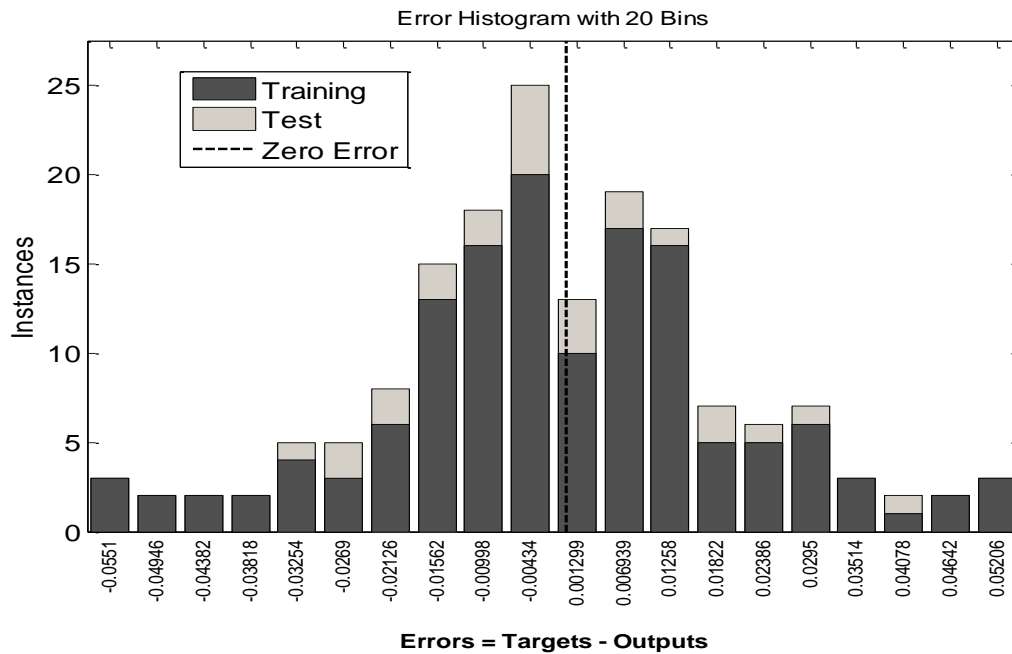


Figure 8: Error analysis on the training, testing and validation data

In addition, evaluation of the established FANN models with the validation or unseen data indicated that FANN models were able to present the prediction equilibrium water dew point of natural gas in TEG dehydration process accurately with high R^2 and low MSE values. In this study, the model has been validated using another 10 data point which is not included in the model development process as shown in Figure 9. This literature data was also act as an unseen data to further evaluate the generalization capability of the developed model. This data's will determined whether the generalization capability of the model developed using training and testing data is acceptable. The regression for the predicted and actual values for

the unseen data is presented in Figure 9. It was found that the FANN model was capable in defining the true behaviour of the process or equilibrium water dew point of natural gas in TEG dehydration process with the R^2 of 0.9972. This data were supported with the analysis of the individual prediction as shown in Table 3. The average absolute deviation percent of error in each sample is 0.5. It implied an excellent generalization capacity of the network or model.

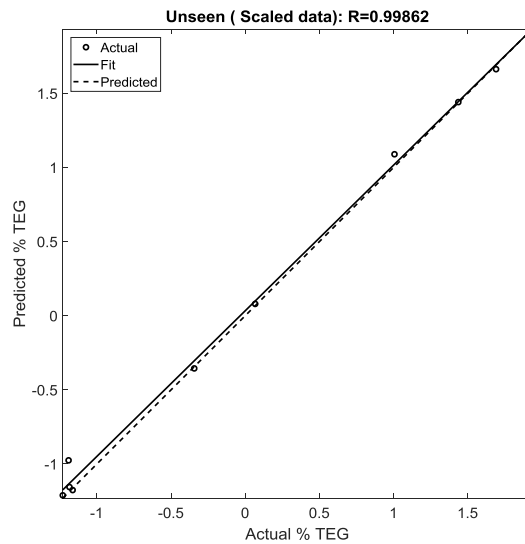


Figure 9: FANN performance on new unseen data

Table 3: Prediction water dew point of a natural gas stream in equilibrium with a TEG solution at various contactor temperatures and TEG concentrations in comparison with the reported data.

TEG weight percent, %	Temperature, K	Predicted dew point (FANN)	Actual dew point from reported data (Gas Processors and Suppliers Association Engineering Data Book, 2004; Parrish et al., 1986; Herskowitz and Gottlieb, 1984)	Absolute deviation, %
90	283.15	267.89	267.15	0.27
95	303.15	278.07	278.15	0.03
97	318.15	284.29	284.65	0.13
98	333.15	290.33	290.65	0.11
99	283.15	243.48	243.15	0.14
99.9	303.15	232.78	232.65	0.06
99.97	283.15	210.67	210.00	0.32
99.99	303.15	210.90	211.70	0.38
99.997	323.15	212.23	211.15	0.51
99.999	343.15	217.54	211.00	3.10
Average absolute deviation percent (AADP)				0.50

By having an accurate model prediction of equilibrium water dew point in natural gas for TEG dehydration process, it will help the researcher to estimate the purity of TEG using empirical model or FANN for this case. It will reduce the time for experimental or developing mechanistic model for the prediction.

Conclusions

In this study, feedforward artificial neural network (FANN) trained by Bayesian regularization method were applied in predicting the equilibrium water dew point in natural gas for TEG dehydration process. Most of the data samples are taken from the literature. The inputs to the network are the contactor temperature and the TEG concentration in percentage, while the output is the equilibrium water dew point. The Bayesian with the regularization training technique was used. Application to the additional unseen data from the literature shows that feedforward artificial neural networks (FANN) with Bayesian training is able to capture and generalize the real process accurately where in this case is equilibrium water dew point in natural gas for TEG dehydration process. The average absolute deviation percentage is very small which is 0.5 % and it is comparable with other studies as well. The presented study indicates that feedforward artificial neural networks (FANN) can effectively model this process.

Acknowledgement

The authors would like to acknowledge the support received from the Universiti Sains Malaysia (USM), Southern Cross University, Lismore NSW Australia and Newcastle University, United Kingdom.

References

- Agatonovic-Kustrin, S., Zecevic, M., Zivanovic, L., Tucker, I.G. (1998) Application of artificial neural networks in HPLC method development. *Journal of Pharmaceutical and Biomedical Analysis*, 17, 69 – 76.
- Bahadori, A., and Vuthaluru, H B. (2009). Rapid estimation of equilibrium water dew point of natural gas in TEG dehydration systems, *Journal of Natural Gas Science and Engineering* 1: 68–71.
- Baş, D., Boyacı, İ.H. (2007) Modelling and optimization II: Comparison of estimation capabilities of response surface methodology with artificial neural networks in a biochemical reaction. *Journal of Food Engineering*, 78, 846 – 854.
- Buntine W L and Weigend A S. 1992. Bayesian Backpropagation. *Complex System* 5: 603-643.
- Chen, J., and Yea, Y. (2002). Neural Network-Based Predictive Control for Multivariable Processes. *Chemical Engineering and Communication* 189: 866-894.

- Coit, D.W., Jackson, B.T., Smith, A.E., 1998. Static neural network process models: considerations and case studies. *International Journal of Production Research*, 36, 2953-2967.
- Desai, K.M., Survase, S.A., Saudagar, P.S., Lele, S.S., Singhal, R.S. (2008) Comparison of artificial neural network (ANN) and response surface methodology (RSM) in fermentation media optimization: Case study of fermentative production of scleroglucan. *Biochemical Engineering Journal*, 41, 266 – 273.
- Gas Processors and Suppliers Association Engineering Data Book, 12th ed., 2004 Gas Processors & Suppliers Association (GPSA), Tulsa, OK, USA.
- Hagan M T, Demuth H B and De Jesus O. (2002). An Introduction to the used of Neural Networks in Control System. *International Journal of Robust and Nonlinear Control* 12: 959-985.
- Herskowitz, M., Gottlieb, M., 1984. Vapour-liquid equilibrium in aqueous solutions of various glycols and polyethylene glycols. *Journal of Chemical & engineering Data* 29, 173.
- Lampinen J and Vehtari A. 2001. Bayesian approach for Neural Network-Review and Case Studies. *Neural Networks* 14: 257-274.
- Lou, W., Nakai, S. (2001) Application of artificial neural networks for predicting the thermal inactivation of bacteria: a combined effect of temperature, pH and water activity. *Food Research International*, 34, 573 – 579.
- McKay D. 1992. A Practical Bayesian Framework for Backpropagation Networks. *Neural Computation* 4: 448-460.
- Parrish, W.R., Won, K.W., Baltatu, M.E. (1986). Paper presented at the 65th GPA Annual Convention, San Antonio, TX, USA.
- Psichogios, D.C., and Ungar, L.H. (1992) A hybrid neural network-first principles approach to process modeling. *AIChE Journal*, 38, 1499 – 1511.
- Radford, M N. (1997). Bayesian Learning for Neural Networks. *Statistics*: 118-183.
- Shaw, A M., Doyle, F J., and Schwaber, J S. (1997). A Dynamic Neural Networks Approach to Nonlinear Process Modelling. *Computers and Chemical Engineering* 21(4): 371-385.
- Shen, C., Wang, L., Li, Q. (2007) Optimization of injection molding process parameters using combination of artificial neural network and genetic algorithm method. *Journal of Materials Processing Technology*, 183, 412 – 418.
- Wang, L., Yang, B., Du, X., Yang, Y., Liu, J., 2008. Optimization of conditions for extraction of acid-soluble collagen from grass carp (*Ctenopharyngodon idella*) by response surface methodology. *Innovative Food Science & Emerging Technologies*, 9, 604-607.
- Xiong, Q., and Jutan, A. (2002). Grey-Box Modelling and Control of Chemical Processes. *Chemical Engineering Science* 57: 1027-1039.
- Zainal, A., Mashitah, M.D., Siti Hatijah, M., Rabiatul Adawiah, M.N. (2010) Nonlinear process modelling of fructosyltransferase (FTase) using bootstrap re-sampling neural network model. *Bioprocess and Biosystem Engineering*, 33, 599 – 606.